بسم الله الرحمن الرحيم

الصفات الواجب توافرها في أدوات جمع البيانات

الصدق Validity

الصدق – مفاهيم أساسية

- يتعلق الصدق بالمدى الذي تقيس فيه أداة معينة ما يفترض أنها تقيسه ويعرف بكونه الصحة والدلالة الهادفة والفائدة للاستدلالات المحددة الناجمة عن درجات الاختبار.
- نحن لا نبرهن على صدق اداة القياس انما نقدم ادلة على تتعلق بصلاحية اداة القياس للغرض الذى استخدمت له.

مثال: قد يكون اختبار مقنن فى القراءة دليل على صدق الاختبار فى تشخيص صعوبات التعلم ولكنه ليس دليلا على صدق الاختبار على على فاعلية برنامج تعليمية معين فى اللغة العربية.

(مثال: اختبار مقنن في التاريخ قد يكون صادقا لمدرس يؤكد على الفهم والتفسير ولكنه غير صادق لمدرس يؤكد على تعلم التواريخ ومعرفة الوقائع)

- لابد من الاهتمام به في العلوم النفسية والاجتماعية لعدم وجود أدوات قياس مباشرة لقياس المفاهيم البنائية.

<u>أنواع أدلة الصدق التي يمكن جمعها:</u>

الصدق المرتبط بالمحتوى Content Related Validity

- يدل على مدى تمثيل اداة القياس للنطاق السلوكي للسمة المراد الاستدلال عليها
- يبين هذا النوع من الدليل المدى الذي تكون فيه عينة بنود الاختبار ممثلة لمجتمع إحصائي محدد أو مجال محتوى . وتجمع الأدلة عن اختبار دقيق وناقد من قبل خبراء بمحتوى الاختبار لتحديد العلاقة بين الاختبار والمجتمع المحدد. وإذا اتفق الجميع على أن بنود الاختبار تمثل مجال المحتوى بشكل مناسب يمكن القول انه يتمتع بصدق المحتوى.
- صدق المحتوى دليل منطقي وليس احصائيا
- يناسب صدق المحتوى الاختبارات التحصيلية (جدول المواصفات)
- صدق المحتوى هو ليس الصدق الظاهري

الصدق المرتبط بالمحك Criterion related validity

يبين الدليل المتعلق بالمعيار المدى الذى ترتبط به درجات اداة القياس مع متغير خارجى مستقل (معيار) يعتقد انه يقيس بشكل مباشر السلوك او الخاصة المعينة (مثال: اختبار الاستعداد والمعدل التراكمى) .

خواص المعيار:

١- وثاقة الصلة بالموضوع (المعيار يمثل الأداء الناجح للسلوك المعنى)
٢- وجوب الوثوق فيه (ينبغي للمعيار أن يكون مقياسا ثابتا للخاصة على مدار الزمن أو من موقف لآخر. فإذا لم يكن المعيار مستقرا فلا يمكن أن نتوقع ارتباطه بشكل ثابت بأي أداة تنبؤ).
٣- أن يكون متحررا من التحيز (أي لا ينبغي أن يتأثر وضع درجات مقياس المعيار بأية عوامل غير الأداء الفعلي الخاص بالمعيار)

المهم الحصول على محكات مناسبة.

أنواعه:

١) الصدق التلازمي (توفر محكات اداء راهنة)

٢) الصدق التنبؤى (توفر اداء محكات مستقبلية)

صدق المفهوم / صدق التكوين الفرضي Construct related validity

- **التكوين الفرضي هو تصور فكرى متعلق بالمفهوم وأبعاده أو هو أفضل ما يتوصل إليه العلماء في وقت ما لتوضيح سمة معينة**
- صدق التكوين الفرضي يركز على وجود مؤشر أو دليل على أن درجات أداة القياس تقيس المفهوم النظري أو السمة المراد قياسها.
- مثال: النضج الاجتماعي لا يمكن قياسه مباشرة، إلا أن العديد من أنواع السلوك التي نعتقد انه تقيس المفهوم يمكن وصفها أو قياسها.

السمة / المفهوم النظرى:	**النضج الاجتماعي: مفهوم نظري لايمكن قياسه بشكل مباشر، يضم عناصر متعددة (مساعدة الذات، التوجية الذاتي، المهنة، الاتصال، العلاقات ..)**
التكوين الفرضى:	التعريف الاجرائي للمفهوم النظري
أداة القياس:	الأداة التي تقيس النضج الاجتماعي حسب التعريف الإجرائي

- وهناك عدة طرق لاستخلاص الأدلة المتعلقة بالمفهوم البنائي وهى:

أولا: أساليب تعتمد على الارتباطات

- دراسة الفروق بين مجموعات مختلفة على مقياس معين ثم المقارنة بين المجموعة العليا والمجموعة الدنيا، وإذا تبين وجود فروق دالة إحصائيا بين المجموعتين يمكن اعتبار هذا احد أدلة صدق التكوين الفرضي للاختبار (صدق المقارنة الطرفية).
- إذا درسنا السمة من خلال استخدام أكثر من أداه/ طريقة، فالارتباطات بين السمات تعتبر دليل على الصدق التقاربي. (مثال: سمة الفهم اللفظي يتم قياسها من خلال: تقديرات المعلمين، اختبارات ورقية، اختبارات فردية)
- إذا درسنا مجموعة من السمات (الفهم اللفظي، الاستدلال العددي، المعلومات العامة) بطريقة واحدة وهى تقديرات المعلمين وكان هناك ارتباطا بين هذه السمات فهذا دليل على الصدق التمايزى.
- التحليل العاملى: هو طريقة إحصائية لدراسة الترابطات المتبادلة بين مجموعة من درجات اختبارات بهدف تحديد عدد العوامل (المفاهيم البنائية) المطلوبة لتفسير هذه الترابطات المتبادلة. فالمرء يبدأ بعدد كبير من المقاييس المختلفة ويفحص الترابطات بينها ويجد العوامل التي تترابط مع بعضها وهذا يعمل على تقليص الفقرات إلى عدد اصغر من العوامل التي يجرى قياسها وتعد أساسا لها.
- صدق الاتساق الداخلي

ثانيا: أساليب تعتمد على التجريب

تعتمد على التدخل التجريبي لإحداث تغيرات في درجات الاختبار كوسيلة للتعرف على مدى تأثر الأفراد بالمعالجات أو متغيرات معينة مما يساعد في تأكيد بعض التفسيرات المتعلقة بنتائج الاختبار أو رفضها.

ثالثا: أساليب تعتمد على التحليل المنطقي

يعتمد على الفحص الدقيق لمحتوى الاختبار والأداء الذي يتطلبه وإحداث تكامل بين نتائج الفحص والنظرية التي يستند إليها الاختبار (التحكيم)

مثال : صدق اختبار قراءة لاغراض مختلفة

نوع الدليل	مثال	اسئلة تطرح
المحتوى	اختبار تحصيل قراءة للصف السادس	مامدى قيام الاختبار بقياس ما جرى تعلمه؟
المتعلق بالمعيار (دراسات متزامنة)	اختبار استعداد للتنبؤ بالاداء فى قراءة الصف السابع	ما مدى قيام الاختبار بالتنبؤ حول تحصيل قراءة الصف السابع؟
المتعلق بالمعيار (دراسات تنبؤية)	اختبار لتحديد اطفال ما قبل المدرسة المعرضين للخطر	ما مدى توافق الاختبار مع ملاحظات اخرى للأطفال؟
المفهوم البنائي	اختبار لقياس استيعاب القراءة	ما مدى قياس اختبار المفهوم البنائي "استيعاب القراءة"؟ هل تستند البيانات والفرضيات حول استيعاب القراءة؟

تدريب: حددي في كل عبارة مما يلي نوع دليل الصدق المناسب:

----------- وجد أن لاختبار ذكاء جديد للاستعداد الدراسي ارتباط مقداره ٠,٩٣ مع اختبار SAT الذي يجرى استخدامه للتنبؤ بالنجاح في الكلية.

----------- لقد تم تطوير اختبار جديد للذكاء ويذهب المؤلف إلى أن العمليات العقلية التي يقتضيها الاختبار تتطابق مع نظرية معينة في الذكاء.

----------- يتفحص مدرس بدقة اختبار تحصيلي مقنن ليرى ما إذا كان يغطى المعرفة والمهارات التي تعلمها الطلبة.

----------- الحكم على اختبار رياضيات من مجموعة مدرسين فيما إذا كانت عينة الأسئلة فيه كافة وممثلة لمجتمع بنود الاختبار.

----------- الحكم على اختبار رياضيات من مجموعة مدرسين فيما إذا كانت عينة الأسئلة فيه كافة وممثلة لمجتمع بنود الاختبار.

الثبات

ما المقصود بمفهوم ثبات درجات الاختبار؟

- مدى خلوها من الاخطاء غير المنظمة التى تشوب القياس اى مدى قياس الاختبار للسمة التى يهدف لقياسها. فدرجات الاختبار تكون ثابتة اذا كان الاختبار يقيس سمة معينة قياسا متسقا فى الظروف المتباينة التى قد تؤدى الى اخطاء القياس.
- الثبات يعنى الاتساق او الدقة فى القياس. وهو معنى بأثر الاخطاء العشوائية للقياس على استقرار الدرجات.

الاخطاء غير المنتظمة (العشوائية) التى ترجع لعوامل الصدفة:

<u>اخطاء عشوائية</u> يصعب التنبؤبها من موقف الى اخر لانها تعمل على خفض الثبات:

- **مصادر متعلقة بالاختبار (اداة القياس)** كعدم وضوح فقرات الاختبار، غموض التعليمات، قلة عدد الاسئلة، عدم ارتباط فقرات الاختبار بالنطاق السلوكى الذى تقيسه (نموذج معاينة النطاق، نموذج الاختبارات المتوازية)
- **مصادرمتعلقة بالظروف البيئية** التى قد تحدث اثناء تقديم الاختبار او باجراءات تطبيق الاختبار وتصحيحه مثل عدم مناسبة وقت الاختبار، الضوضاء وقت الاختبار، التهوية، الاضاءة ، تعليمات الاختبار غير واضحة وبخاصة فى الاختبارات الفردية الشفوية التى تطبق على الاطفال او الاميين، كيفية تصحيح الاختباروخاصة للاسئلة المفتوحة ...
- **مصادر متعلقة بالافراد المختبرين** مثل شعورهم بالتعب وقلة الدافعية ، المرض، القلق، عدم امتلاكهم لمهارات التعامل مع الاختبار، عدم القدرة على القراءة، تذكر الاجابات عن مفردات الاختبار نتيجة الخبرة السابقة او اعادة تطبيق الاختبار..

الأخطاء المنتظمة:

يقصد بها الاخطاء الدائمة الثابتة . ... لها علاقة بصدق الاختبار حيث انه قد يكون الاختبار يقيس بشكل متسق سمة غير السمة التي اعد لقياسها. مثلا مقياس اعد لقياس الانبساطية يقيس أيضا عن غير قصد القلق بشكل متسق.

الثبات والدرجات الحقيقية في النظرية الكلاسيكية في القياس (الدرجة الحقيقية والدرجة الخطأ)

طور النظرية الكلاسيكية شارل سبيرمان ١٩٠٤ ثم تم تطويرها من قبل عدد من السكومتيرين المعاصرين. والنظرية ما زالت سائدة رغم ظهور نماذج حديثة فى القياس. ومنطلق النظرية أن درجة الاختبار تنتج عن تأثير عاملين هما:

١- العوامل التي تسهم في الاتساق (الخصائص المستقرة للفرد والتي يحاول الباحث قياسها)

٢- العوامل التي تسهم في عدم الاتساق (الأخطاء الناتجة عن القياس للأفراد، تطبيق الاختبار، اختيار المفردات، تصحيح الاختبار ..)

اى ان الدرجة التي يحصل عليها الفرد فى الاختبار ليست هى الدرجة الحقيقية انما هى الدرجة الحقيقة مضافا اليها درجة الاخطاء العشوائية.

- الدرجة الحقيقة = الدرجة الملاحظة - درجة الاخطاء العشوائية
- الدرجة الحقيقة لا يمكن معرفتها على الإطلاق وكل ما يمكن ان نصل اليه هو تحديد احتمالية ان تقع هذه الدرجة الحقيقة فى مدى معين . وانه من الممكن ان نحصل على احسن تقدير للدرجة الحقيقية ولكن من المستحيل معرفة قيمتها.
- اذا كانت درجات الاختبار غير متسقة معنى ذلك ان الدرجات لاتمثل سوى اخطاء قياس اي ان الفروق الملاحظة بين الافراد فى هذه الحالة لا يكون لها علاقة بالفروق الحقيقة بينهم وتكون قيمة معامل الثبات صفرا.
- خطأ القياس هو كل شيء يمكن ان نحصل عليه بخلاف الدرجة الحقيقية. ومصادر الخطأ تتضمن: اختيار المفردات، تطبيق الاختبار، تصحيح الاختبار، والاخطاء المنتظمة للقياس. وخطأ القياس يقلل من ثبات الاختبار.
- الانحراف المعيارى لدرجات الخطأ يسمى الخطأ المعيارى للقياس.
- سؤال: ماذا لو كان الخطأ المعياري للقياس = صفرا .. هذا معناه أن درجات الاختبار فيها اتساقا كاملا.
- ملاحظة: في العلوم النفسية التربوية عادة ما يكون هناك قدر من الخطأ المعيارى وهذا يؤدى الى خفض قيمة الثبات.

النظرية الكلاسيكية

الخصائص الأساسية (الافتراضات الأساسية) للنظرية الكلاسيكية

١. **أخطاء القياس عشوائية (غير منتظمة).** اى لا يمكن التهكن باثارها والاثار غير متسقة.
٢. **متوسط خطأ القياس = صفر .** مثلا: خطا قياس ناتج عن التشويش السمعى لشخص يسعل اثناء تطبيق الاختبار، شخص اخر استفاد من خطا قياس موجب عن طريق استرقاق النظر فى المراة الموجودة خلف الفاحص ليرى الاجابة الصحيحة لبند معين في اختبار معلومات.
٣. **الدرجات الحقيقية ودرجة الخطأ ليس بينما ارتباط .** اخطاء القياس لا ترتبط مع الدرجة الحقيقية، واذا كانت مرتبطة بدرجة اخرى كان من الممكن ان توحى بانها تكون منتظمة اكثر من كونها عشوائية
٤. **الأخطاء على المقاييس المختلفة ليس بينهما ارتباط أي أنها مستقلة عن بعضها البعض**

من هنا يمكننا ان نطور استخدامات عامة للقياس والثبات. فاذا طبقنا مقياس مع مجموعة كبيرة من المفحوصين فيسيظهر لنا تباينا كبرا فى الدرجات المشاهدة ويمكن التعبير عن ذلك فى صورة التباين.

واهمية النظرية انها تقسم التباين بالدرحات المشاهدة الى مصدرين منفصلين هما: تباين الدرجات الحقيقة وتباين اخطاء درجات القياس. المعادلة هى:

تباين الدرجات المشاهدة = تباين الدرجة الحقيقية + تباين اخطاء القياس

مما سيق نرى ان الثبات يعبر عن التأثير النسبى للدرجات الحقيقية ودرجات الخطأ على درجات الاختبار المشاهدة

معامل الثبات

- معامل الثبات = تباين الدرجات الحقيقية / تباين الدرجات الملحوظة
- تتراوح قيم معامل الثبات ما بين ٠ و ١
- المقياس الذى فيه يتمتع بدرجة ثبات عالية هو الذى تقل فيه اخطاء القياس الى ابعد حد

مؤشر الثبات

- العلاقة بين الدرجات الحقيقية (غير المعلومة) والدرجات الملاحظة (المعلومة)
- معامل الثبات هذا يدل على العلاقة ما بين الدرجات الملحوظة بين الاختبارين.
- مثال: قام باحث بتطبيق نفس الاختبار على مجموعة من الافراد . قام بحساب معامل الثبات وكان = ٠,٨٥
- اذا اردنا حساب مؤشر الثبات نقوم بايجاد الجذر التربيعى لمعامل ثبات درجات الاختبار.
- مؤشر الثبات فى المثال السابق = ٠,٨٩

Types of Reliability

أنواع أدلة الثبات التي يمكن جمعها

من خلال اعادة الاختبا ر Stability

- نقوم بحساب معامل الاستقرارمن خلال تطبيق الاختبار مجموعة معينة ثم اعادة تطبيقة على نفس المجموعة بعد مرور فترة زمنية معينة ثم حساب معامل ارتباط بيرسون.
- المصدر الاساس للاخطاء العشوائية قد يكون بسبب تذكر الافراد لفقرات الاختبار او حصول تغيرات على الأفراد انفسهم بسبب الزمن ..

Equivalence الصور المتكافئة

- نقوم بحساب معامل التكافؤ من خلال تطبيق صيغتين متكافئتين من الاختبار. ثم حساب معامل ارتباط بيرسون.
- المصدر الاساس للاخطاء العشوائية هو الفروق بين مفردات كل من صيغتى الاختبار او المقياس التى ربما تكون نتيجة عدم تمثيل مفردات كل من الصيغيتن للنطاق الشامل تمثيلا كافيا.

Equivalence and stability من خلال اعادة الاختبار والصور المتكافئة

هذه الطريقة تجمع بين الطريقيتن السابقتين وعادة ما تؤدى الى قيم تقديريه لمعامل الثبات اقل من نظريتها فى اي من الطريقيتن السابقتين وذلك لانها تجمع بين الاخطاء العشوائية التى تؤثر فى كل منهما

التوافق الداخلى/ الاتساق الداخلى Internal consistency

تطبيق الاختبار مرة واحدة واجراء معامل الارتباط بين نصفى الاختبار (لا بد من التأكد من تجانس نصفى الاختبار ولا يجوز استخدامه فى اختبارات السرعة).

طريقة التجزئة النصفية Split-half

يتم حساب معامل الارتباط بين تصفى الاختبار ثم عمل تعديل باستخدام معادلة سبيرمان براون (الافتراض عند استخدام هذه المعادلة تساوى تباين الدرجات على نصفى الاختبار – اى تجانس التباين وهذا شرط من شروط التكافؤ)

اذا لم يتحقق هذا الافتراض تستخدم معادلة بديلة تاخذ بالاعتبار اختلاف التباين اختلافا جوهريا وهى معادلة **جوتمان**

<u>طريقة كرونباخ والمعادلات المشتقة منها Cronbach alpha</u>

(تعتمد على حساب الارتباطات بين العلامات لمجموعة الثبات على جميع الفقرات الداخلة فى الاختبار، وكاننا قسمنا الاختبار الى عدد من الاجزاء يساوى عدد فقراته فمثلا اذا كان عدد الفقرات = ٢٠ فإن عدد الارتباطات = ٢٠* ١٩/٢ = ١٩٠ ارتباط)

معادلة كرونباخ - تستخدم هنا اى نوع من انواع الاسئلة الموضوعية او المقالية

معادلة كودرريتشاردسون Kuder-Richardson 20

يعتمد على نسبة الاجابات الصحيحة وغير الصحيحه لكل بند من الاختبار

تستخدم هنا اى نوع من انواع الاسئلة الموضوعية (٠، ١)

معادلة كودرريتشاردسون Kuder-Richardson 21

تفترض ان صعوبة الفقرات متماثلة

<u>ثبات الاتفاق بين المحكمين Agreement</u>

يعتمد على ايجاد فيمة معامل الارتباط بين تقديرى محكمين مختلفين او درجة الاتفاق بين الملاحظين (معامل كابا)

<u>التحليل العاملى</u> وجود عامل واحد يجمع بين الفقرات مؤشر للثبات

<u>تحليل التباين صيغة هويت</u>

Procedures for estimating reliability

	Time 1	**Time 2**
Stability	A	A
Equivalence	AB	
Equivalence and stability	A	B
Internal consistency	A	
Agreement	R1 R2	

* A and B refer to different forms of the same test, R1 and R2 refer to rater 1 and rater

ما هو الثبات المناسب ؟

من الشائع ان تشتمل ادلة الاختبارات على عرض لمصادر متعددة للمعلومات الخاصة بالثبات

تفسير معامل الثبات

- يعتمد ثبات الاختبار جزئيا على طول الاختبار
- يعتمد الثبات على عدم تجانس المجموعة، فمعامل الثبات يزداد عندما يزداد انتشارا او عدم تجانس الاشخاص الذين يأخذون الاختبار.
- يتاثر معامل بزمن الاختبار (يرتفع بزيادة الزمن حتى حد معين)
- يعتمد الثبات على قدرة الاشخاص الذين يأخذون الاختبار. فقد يكون الاختبار ثابتا على مستوى قدرة واحد ولكنه لا يكون ثابت على مستوى اخر.
- يعتمد الثبات على الاسلوب الخاص المستخدم لتقديره
- يعتمد الثبات على طبيعة المتغير المقاس

ما هو الحد الأدنى لثبات مقبول لاداة معينة؟

- هذا يعتمد على السمة المقاسة وعلى استخدام النتائج المترتبة وبالذات بالنسبة للاشخاص.
- المستوى المقبول يتوقف على كمية خطأ القياس التى يمكن لمستخدم الاختبار ان يتحمله فى التطبيق المقترح للاختبار.

تدريب : حددي في كل عبارة مما يلي نوع معامل الثبات المناسب:

----------- قام مدرس بعمل صيغتين متكافئتين للاختبار، طبق الصيغة الأولى قبل البدء في عملية التعلم والأخرى بعد الانتهاء من عملية التعلم

----------- أخصائي اجتماعي طور أداة لقياس اتجاهات الطلبة الجدد نحو الحياة الجامعية واستخدم المقياس التالي: ٢= نعم ١=ربما ٠=لا

----------- قام أخصائي نفسي بتطبيق صيغتين متكافئتين (أ) و (ب) من اختبار نفسي على مجموعة مكونة من ٢٠ فرد. فالصورة (أ) طبقت في يوم ما يليها طبقت الصورة (ب) في نفس اليوم.

----------- استخدم مدرس نسخة وحيدة من اختبار تحصيلي مقنن متعلق بالعلوم الصحية واستخدم المقياس التالي ٠، ١ وكانت موضوعات الاختبار هي: التغذية، التغلب على المرض، الصحة النفسية. وحرص المدرس أن تكون درجات صعوبة الأسئلة متجانسة.

----------- المدرسة أمل أرادت أن تفحص معامل الثبات لاختبار تحصيلي فقراته غير متساوية الصعوبة في مادة الرياضيات المكون من أسئلة اختيار من متعدد.

----------- باحث طور اختبار لقياس اتجاه الأفراد نحو استخدام الحاسوب. استخدم الاختبار قبل البدء في البرنامج وأعاد استخدامه بعد البدء في البرنامج في ظروف مناسبة ومحددة.

----------- قامت الباحثة هديل بإيجاد معامل الارتباط بين درجات كل من نصفي الاختبار بعد الانتهاء من التطبيق ثم قامت بإجراء تعديل على قيمة معامل الارتباط.

أهم الأساليب والأدوات المستخدمة في جمع المعلومات في البحوث التربوية المختلفة:

الاختبارات التحصيلية Achievement tests

من المهم أن تتمتع أدوات القياس والتقويم الصفي بالصدق Validity

ويقصد بالصدق قياس أدوات القياس لما أُعدت لقياسه. وهو مرتبط بالاستدلالات والنتائج التي نصل إليها بعد عملية القياس والتقويم، وهل هي مرتبطة بالغرض الذي تمت لأجله عملية القياس والتقويم؟ فمثلا: إذا صممت أداة لقياس قدرة طلبة المستوى السادس على استخدام الحاسوب ، فيجب أن تقيس الأداة هذه القدرة التي صممت لأجلها، أما إذا قاست اتجاهات الطلبة نحو الحاسوب فهي أداه غير صادقه .

من المهم أن نلاحظ أنه لا يمكننا أن نبرهن على صدق الاختبار وإنما نستطيع أن نقدم أدلة تتعلق بصلاحية استخدامه في غرض أو أغراض معينة.

اقتراحات لتعزيز صدق أدوات القياس والتقويم

اطلب من زملائك المعلمين الحكم على وضوح ما تقوم بقياسه من الأدوات التي تستخدمها.

- تأكد من أن الطرق المختلفة المستخدمة لقياس نفس الشيء تعطى نفس النتائج.
- خذ عينات كافية من الأمثلة لما يتم قياسه.
- قم بإعداد جدول مواصفات .
- اطلب من الآخرين الحكم على مدى المطابقة بين بنود التقويم وأهداف عملية التقويم.
- استخدم أدوات قياس مختلفة لقياس نفس الأشياء.
- تأكد من سلامة اللغة ومدى صعوبة الفقرات.
- ابدأ بالأسئلة السهلة في الاختبارات.
- أعط وقت كاف لتطبيق أدوات القياس والتقويم.
- استخدام أدوات القياس لقياس مدى تحقق الأهداف التدريسية فقط.
- قارن بين مجموعات معروفه باختلافها على الشيء الذي يتم قياسه.
- قارن بين الدرجات قبل عملية التدريس ثم بعدها.
- قارن بين النتائج المتوقعة بالنتائج التي حصلت عليها.

من المهم أن تتمتع أدوات القياس والتقويم الصفي بالثبات Reliability.

ويقصد بالثبات أن مركز الطالب النسبي لا يتغير إذا أعيد الاختبار على الطالب نفسه . وهذا يعني استقرار النتائج عند تكرار تطبيق الاختبار أو صور مكافئة له على المجموعة نفسها من الأفراد .

اقتراحات لتعزيز ثبات أدوات القياس والتقويم :

- استخدم عدد كاف من الأسئلة في المهام التقويمية (كلما كان عدد الأسئلة أكثر كلما كانت الأداة التقويمية أكثر ثباتا).
- استخدم مقيمين مستقلين للحصول على درجات متشابهه عند تقويم المهام الأدائية.
- طور أسئلة ومهام تقويمية تميز بين الطلبة.
- تأكد من أن إجراءات عملية التقويم موضوعية.
- استمر في عملية التقويم حتى تحصل على نتائج متسقة.
- قلل من تأثير العوامل الخارجية على عملية التقويم ما أمكن.
- استخدم المهام التقويمية القصيرة بشكل مستمر بدلا من استخدام مهام تقويمية طويلة وعددها قليل.

إعداد جدول المواصفات

بعد تحديد محتويات المقرر الدراسي والأهداف التعليمية يقوم المعلم بتحضير جدول المواصفات. **ثم يعرف جدول المواصفات** بأنه مخطط تفصيلي يربط العناصر الأساسية للمحتوى بالأهداف التدريسية بمستوياتها المختلفة. أو هو عبارة عن مخطط تفصيلي للاختبار يتحدد من خلاله عدد الأسئلة الممثلة للمحتوى والمحققة لمستويات الأهداف. ويتكون من بعدين: رأسي: ويمثل مستويات الأهداف ، والثاني: أفقي يمثل موضوعات محتوى المادة. وكل خلية من خلايا الجدول تشير إلى ارتباط محتوى معين بهدف ذي مستوى معين.

وخطوات إعداده تتمثل فيما يلي:

١-تحديد النواتج وعناصر المحتوى الذي سيجرى قياسها.

وهنا يبدأ المعلم بتقسيم المادة الدراسية إلى موضوعات أو عناوين رئيسه، كما يمكن تقسيم الموضوعات الرئيسة إلى عناوين فرعية (ليس بالضرورة للمعلم أن يتقيد بالعناوين الواردة في الكتاب المقرر إذا كانت كثيرة العدد). من المهم أن يدرك المعلم أن زيادة الموضوعات والعناوين تعنى ضمنا زيادة عدد الأسئلة في الاختبار.

وبعد تحديد الموضوعات يقوم المعلم بتحديد وزن وأهمية كل موضوع بالنسبة للموضوعات الأخرى وذلك من خلال الرجوع إلى الزمن الذي استغرقه في تدريس الموضوع، أو إلى عدد الصفحات التي يشغلها الموضوع في الكتاب المقرر، أو إلى مدى مساهمة الموضوع في تعلم لاحق. وتساعد الأهداف التدريسية التي وضعها المعلم في تحديد النسب المئوية أو الأوزان النسبية للموضوعات. ولنفترض مثلا أن الأهداف التدريسية التي يقيسها الاختبار مركزة في (٣) موضوعات هي: الوحدة الأولى، والوحدة الثانية، والوحدة الثالثة بنسب هي ٢٠%، ٣٠%، ٥٠%.

٢-تحديد المجالات والمستويات التي يريد المعلم قياسها[١].

وهنا في هذا الفصل سنركز على إعداد جدول مواصفات يركز على المجال المعرفي ومستوياته المختلفة؛ لأن اغلب الاختبارات التحصيلية تركز على الأهداف بالمجال المعرفي؛ نظرا لصعوبة قياس الأهداف في المجالين النفس حركي والانفعالي. وخاصة وان الأهداف النفس حركية تركز على الأداء وقد تقاس من خلال الملاحظة واختبارات الأداء. وبعد تحديد المستويات المعرفية يقوم المعلم بتحديد وزن وأهمية كل مستوى معرفي. لنفترض مثلا أن الأهداف المعرفية التي سيقيسها الاختبار هي في مستويات الفهم، والتطبيق، والمستويات العليا). وأن الأوزان التي تناسب عدد وأهمية الأهداف في كل مستوى هي على التوالي (٣٠%، ٣٠%، ٤٠%).

٣- تحديد عدد فقرات الاختبار آخذا بالاعتبار العوامل المحددة لطوله فإذا كان العدد المناسب للفقرات ١٠٠ فقرة من نوع اختبار من متعدد مثلا فانه يمكن تحديد عدد الفقرات لكل موضوع (مستوى) . مثلا عدد الفقرات في الوحدة الأولى في مستوى الفهم تساوى (٢٠% x ٣٠%) / ١٠٠ = ٦ فقرات

ويمكن تلخيص الخطوات السابقة في جدول ذي بعدين بحيث يبين أحد البعدين الموضوعات والنسب المحددة لأوزانها في حين يبين البعد الثاني المستويات وأوزانها كما يبين عدد الفقرات الممثلة لتقاطع كل موضوع (مستوى) في كل خلية من خلايا الجدول. ويبين الجدول التالي جدول المواصفات للمثال الوارد في الخطوات السابقة.

جدول مواصفات

الموضوع	المستوى المعرفي			المجموع
	فهم %٣٠	تطبيق %٣٠	مستويات عليا %٤٠	%١٠٠
الوحدة الأولى ٢٠%	٦ فقرات	٦ فقرات	٨ فقرات	٢٠
الوحدة الثانية ٣٠%	٩ فقرات	٩ فقرات	١٢ فقرة	٣٠
الوحدة الثالثة ٥٠%	١٥ فقرة	١٥ فقرة	٢٠ فقرة	٥٠
المجموع ١٠٠%	٣٠	٣٠	٤٠	١٠٠

يمكن أن يوفر الجدول درجة مقبولة من صدق محتوى الاختبار، أو صدق تمثيل عينة الفقرات للأهداف إذا تم تقسيم الموضوعات وأوزانها والمستويات وأوزانها على أسس منطقية يدافع عنها المعلم الذي وضع الاختبار.

[١] يمكن أن يستخدم المعلم جدول المواصفات للقياس في المجالات المختلفة (المجال المعرفي، المجال النفس حركي، المجال الانفعالي)

من الملاحظ أن العدد الكلى لفقرات الاختبار يتحكم في عدد الفقرات داخل كل خلية من خلايا الجدول، كما أن العدد يتغير بتغير عدد الخلايا مع ثبات العدد الكلي.

مثال تطبيقي: إعداد جدول مواصفات في مبحث الجغرافيا (الصف السابع الأساسي).

الأهداف التعليمية	الوحدة	مستوى الهدف المعرفي
١- أن يعّرف الموقع الجغرافي	١	تذكر
٢- أن يفسر أهمية دراسة الموقع الفلكي	١	فهم
٣- أن يميز بين الموقع الفلكي والموقع الجغرافي	١	فهم
٤- أن يرسم خريطة فلسطين مبينا عليها أسماء المدن، البحار، الأنهار	١	تطبيق
٥- أن ينتقد خريطة مصممة حديثا	١	تقويم
٦- أن يبين رأيه في تضاريس فلسطين	١	تقويم
٧- أن يعّرف الصخور النارية	٢	تذكر
٨- أن يوضح أشكال التعرية بالرياح	٢	فهم
٩- أن يقارن بين العوامل الباطنية والعوامل الخارجية المؤثرة في تشكيل سطح الأرض	٢	تحليل
١٠-أن يبين مميزات سهل مرج ابن عامر	٢	فهم
موضوعات الاختبار:		
الوحدة رقم ١	٦	%٦٠
الوحدة رقم ٢	٤	%٤٠
المستويات المعرفية (العقلية) لأهداف الاختبار		
تذكر	٢	%٢٠
فهم	٤	%٤٠
تطبيق	١	%١٠
مهارات عليا	٣	%٣٠

جدول مواصفات

	تذكر (%٢٠)	فهم (%٤٠)	تطبيق (%١٠)	مهارات عليا (%٣٠)	المجموع
الوحدة ١ (%٦٠)	١٢	٢٤	٦	١٨	٦٠
الوحدة ٢ (%٤٠)	٨	١٦	٤	١٢	٤٠
المجموع	٢٠	٤٠	١٠	٣٠	١٠٠